



# Alternative Control Groups

When Placebo Doesn't Match Your Trial Design

---

David Snead, Ph.D.

Principal Biostatistician

Cordis Corporation, a Johnson &  
Johnson Company



# Outline

---

- Critical Path opportunities
- Types of control groups ICH E 10
  - Active controls
  - Historical controls
- Statistical adjustments
- Innovative control groups
  - Virtual controls
- Questions and answers



# FDA's Critical Path white paper

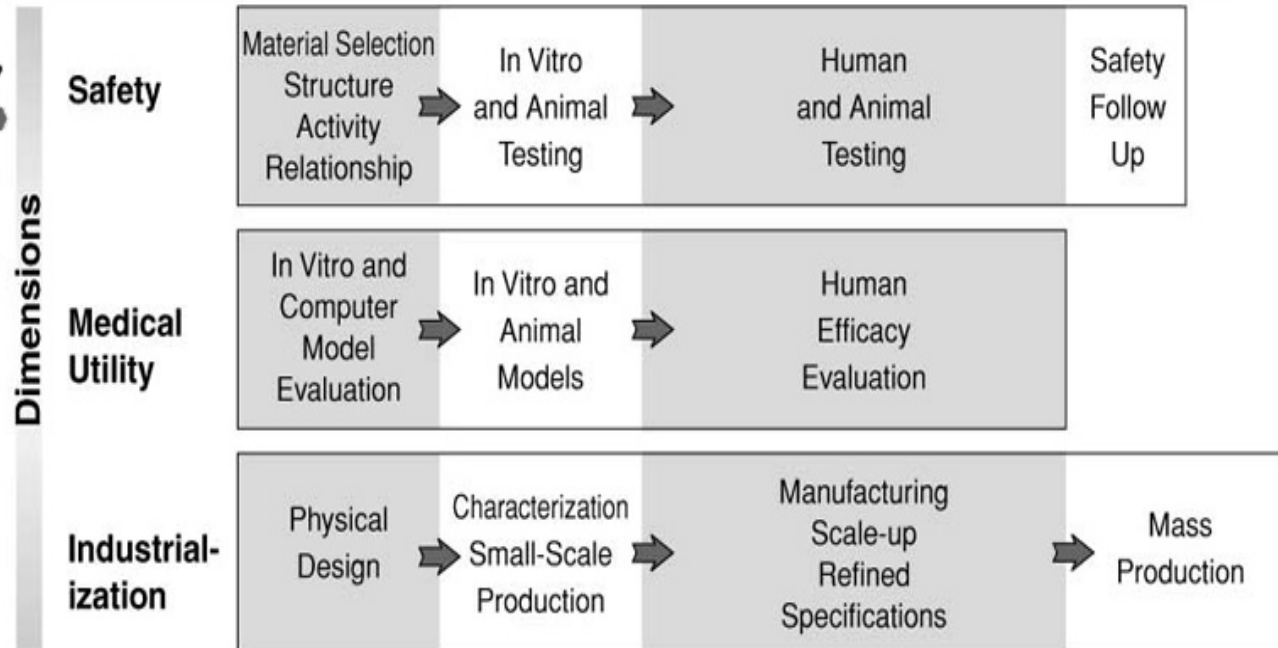
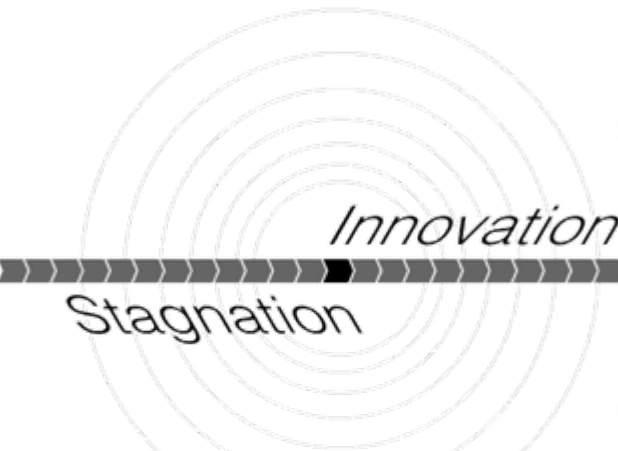
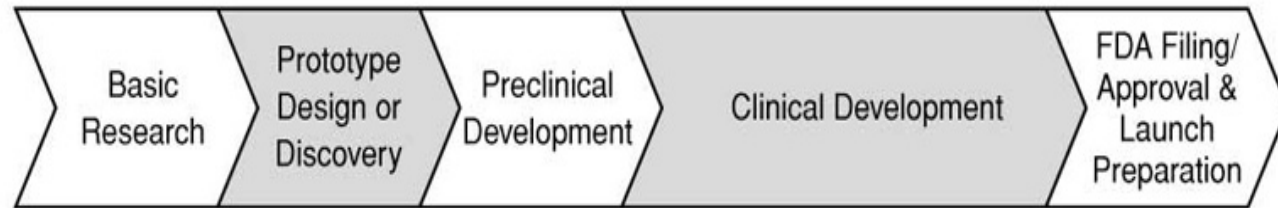
---

- Analyzes the hurdles of medical product development – *can't keep pace with basic scientific innovation*
- Calls for collaboration between government, research institutions and manufacturers to promote the public health by getting safe and effective medical products out to the public as quickly as possible



FDA SCIENCE:  
**THE CRITICAL PATH**  
 FROM CONCEPT  TO CONSUMER

May 18-19, 2004 • [www.dcscienceforum.org](http://www.dcscienceforum.org)



**Challenge and Opportunity  
 on the Critical Path  
 to New Medical  
 Products**

# Critical Path research



Leverages basic science knowledge

Leverages cumulative research experiences

*Does not compromise safety and effectiveness evaluations*



## Why FDA?

- FDA has the broad perspective regarding why some products fail and others succeed
- Companies lack this information
- Academic centers and Agencies such as NIH are not faced with problems of product development and manufacture



# Critical Path Opportunity List

---

- 76 opportunities identified in the areas of *evaluative tools, streamlining clinical trials, harvesting bioinformatics*, manufacturing, urgent public needs and pediatrics
- The Critical Path to New Medical Products available at:  
<http://www.fda.gov/oc/initiatives/criticalpath/>



# Critical Path Opportunities: *Advancing innovative trial designs*

---

- 34. Design of Active Controlled Trials
- 35. Enrichment Designs
- 36. Use of Prior Experience or Accumulated Information in Trial Design
  - 1. Adaptive Trial Design
  - 2. Non-Frequentist Methods – Bayesian
- 37. Development of Best Practices for Handling Missing Data
- 38. Analysis of Multiple Endpoints



# Critical Path Opportunities: *Harnessing bioinformatics*

---

- 46. Identification and Quantification of Safety Biomarkers
- 47. **Virtual Control Groups in Clinical Trials**
- 48. Adverse Event Data Mining
- 49. Multiple Complex Therapies
- 50. Modeling Device Performance
- 51. **Clinical Trial Simulation**
- 52. Failure Analysis
- 53. **Natural History Databases for Rare Diseases**



# Control group

## Guidance for industry

---

- E 10 Choice of Control Group and Related Issues in Clinical Trials  
FDA (CDER & CBER) ICH May 2001  
available at:

<http://www.fda.gov/cder/guidance/4155fnl.htm>

- All FDA and ICH guidance docs  
available at:

<http://www.fda.gov/cder/guidance/index.htm>



# Purpose of control group\*

---

- Allows discrimination of patient outcomes caused by the test treatment from outcomes caused by other factors, such as the natural progression of the disease, observer or patient expectations, or other treatment
- The control group experience tells us what would have happened to patients if they had not received the test treatment or if they had received a different treatment known to be effective

\*E 10 Choice of Control Group and Related Issues in Clinical Trials



# Types of controls

---

- Classified on two attributes\*
  - Type of treatment used
  - Method of determining who will be in control group
    - Randomization
    - Selection (external or historical)

\*E 10 Choice of Control Group and Related Issues in Clinical Trials



# Types of control groups\*

---

- Placebo concurrent control
- No treatment concurrent control
- Dose-response concurrent control
- Active (positive) concurrent control
- External control (including historical control)
- Multiple control groups

\*E 10 Choice of Control Group and Related Issues in Clinical Trials

# Usefulness of concurrent control types (1)\*

	Control Type			
Trial Objective	Placebo	Active non-inferiority	Active superiority	Dose Response (D/R)
Measure <i>Absolute</i> effect size	Y	N	N	N
Show existence of effect	Y	P	Y	Y
Show Dose-Response relationship	N	N	N	Y
Compare therapies	N	P	Y	N

Y=Yes, N=No, P=Possible, depending on whether there is historical evidence of sensitivity to drug effects

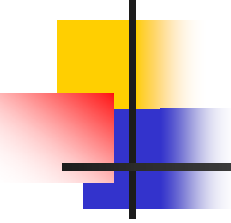
\*E 10 Choice of Control Group and Related Issues in Clinical Trials

# Usefulness of concurrent control types (2)\*

	Control Type			
Trial Objective	Placebo + Active	Placebo + D/R	Active + D/R	Placebo + Active + D/R
Measure <i>Absolute</i> effect size	Y	Y	N	Y
Show existence of effect	Y	Y	Y	Y
Show Dose-Response relationship	N	Y	Y	Y
Compare therapies	Y	N	P	Y

Y=Yes, N=No, P=Possible, depending on whether there is historical evidence of sensitivity to drug effects

\*E 10 Choice of Control Group and Related Issues in Clinical Trials



Keeping with theme of the session: control types\*  
When placebo doesn't match your trial design

---

- Placebo concurrent control
- *No treatment concurrent control*
- *Dose-response concurrent control*
- *Active (positive) concurrent control*
- *External control (including historical control)*
- *Multiple control groups*

\*E 10 Choice of Control Group and Related Issues in Clinical Trials



# Placebo concurrent control

---

- We all know the value of placebo control
  - Controls for the well documented “placebo effect”
  - Controls for all potential influences on course of disease other than test therapy
  - Could be
    - **No treatment + placebo**
    - **Standard care + placebo**



# Other concurrent control types

---

- No treatment concurrent control – essentially same as placebo but unable to blind → *use blinded reviewers*
- Dose response concurrent control – several variations with or without placebo → *not covering here*
- Multiple control groups – gets complex → *not covering here*



# Active concurrent controls

---

- Random assignment to test or active control treatment
- 2 distinct objectives:
  - Show superiority of test to active control – well understood methods → *not covering*
  - Show test is as good as known effective treatment – non-inferiority/equivalence → *not so well understood or carried out so let's look at*



# Equivalence/non-inferiority trial (1)

---

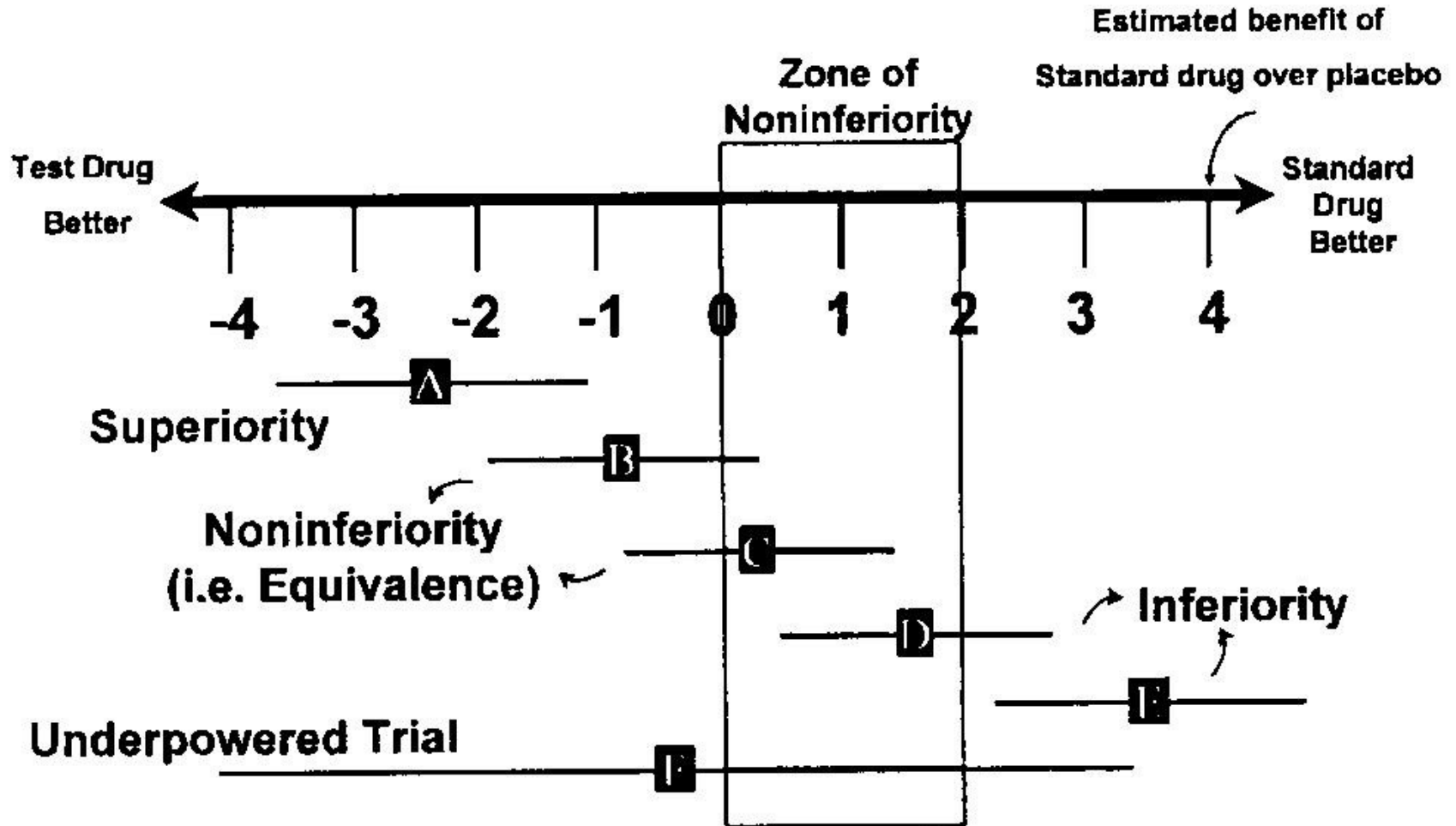
- Late 1960's concept of bioavailability and bioequivalence became public issue with generic drugs → too little drug ↓ efficacy and too much drug ↑ safety concerns
- FDA began requiring evidence in pharma studies leading to equivalence statistical methods in 70-80's and still evolving today
- The question is whether new (easier or cheaper) treatment is as good as the current treatment
- Must specify margin of "equivalence" or non-inferiority (NI)
- Can't statistically prove equivalency → only show that difference is less than something with specified probability



## Equivalence/non-inferiority trial (2)

- Need historical evidence of sensitivity to treatment → control clearly better than placebo
- Sample size issues are crucial
- Small sample size, leading to low power and subsequently lack of significant difference, does not imply “equivalence”
- Must use NI hypothesis and statistical test
  - $H_0: \mu_T - \mu_R \leq \delta_L \text{ or } \mu_T - \mu_R \geq \delta_U$
  - $H_a: \delta_L < \mu_T - \mu_R < \delta_U$
- Testing for difference and finding no significant difference ( $P > 0.05$ ) does not mean 2 things are equivalent, although this was commonly done in the past

# Difference in events test drug – standard drug





# Avoid “bio-creep” with NI

---

- Is the control an appropriate one? Compare to best licensed product, not worst
- Never use a generic drug comparator as long as innovator drug is available because could lead to bio-creep phenomenon, resulting in progressively less reliable similarity of future multi-source products and to lack of interchangeability with the innovator
- Problem with repeated NI trials



# Guidelines for establishing deltas

---

- ICH E10: “The determination of the margin in a NI trial is based on both statistical reasoning & clinical judgment, should reflect uncertainties in the evidence on which the choice is based, and should be suitably conservative.”



# Guidelines for establishing deltas

## “Step-Function” Delta in Antimicrobial

---

Delta in NI trials (as described in FDA's 1992 Points to Consider document) has been useful for assessing efficacy, using reasonable patient sample sizes, especially when multiple Phase 3 trials are done

<b><u>Cure Rate</u></b>	<b><u>Delta</u></b>
<b>≥ 90%</b>	<b>10%</b>
<b>80 - 89 %</b>	<b>15%</b>
<b>&lt; 80 %</b>	<b>20%</b>



# Guidelines for establishing deltas

---

- Use an existing delta – don't re-invent wheel
- What is clinically important? Depends on context of disease – for clinical event endpoint may use 25-50% of expected failure rate
  - get consensus from clinical experts
- Bioequivalence often 20% blood levels of drug
- Drug content often 10% amount of drug
- Examine the variance in the data
- Get agreement from customer (e.g., FDA)



# Randomized Controlled Trial (RCT) is gold standard

---

- No doubt, whether control is placebo or standard of care or different dose randomization is the best way to allocate treatment assignment
- Situations where randomization is not feasible or ethical



# Disadvantages of RCT

---

1. Generalizable results?
  - subjects may not represent general patient population – volunteer effect
2. Recruitment
  - twice as many new patients
3. Acceptability of randomization process
  - some physicians will refuse
  - some patients will refuse
4. Administrative complexity
5. Unethical in some circumstances



# External control

---

- Compares group on test treatment (prospectively) with group external to the trial
  - Concurrent, non-randomized control (not really external to trial)
  - non-concurrent, non-randomized control = historical control
- Should be a defined group of patients, preferably with same I/E criteria, same study procedures followed, etc.
- Need study hypothesis and sample size justification
- Biggest concern → large differences on covariates could lead to biased estimates of treatment effect



# Concurrent controls

## *Not in E 10 Guidance?*

---

- Not randomized because not feasible
- Patients treated by different strategies during same time period (and sites) and compared, e.g., new treatment vs. standard of care
- Advantages
  - Eliminate time trend differences (seen with historical control)
  - Data should be of comparable quality
- Disadvantages (same as historical control)
  - Selection bias
  - Treatment groups may not be comparable



# Biases in concurrent control

---

- Sources of bias
  - Patient selection
    - referral patterns
    - refusals
    - different eligibility criteria
  - Experimental environment
    - diagnosis/staging
    - supportive care
    - evaluation methods
    - data quality
- Do everything possible to minimize these by keeping everything the same in both arms!



# Historical control (1)

---

- A new treatment prospectively used in a series of patients preferably with same I/E criteria, same study procedures followed, etc. as historical control
- Compared with previous series of comparable patients, e.g., earlier study for same indication
- Non-randomized, non-concurrent control
- Rapid, inexpensive, good for initial testing of new treatments or when can't randomize
- Two sources of historical control data:
  - Literature – subject to publication bias → not covering as usually unacceptable to FDA
  - Database – patient-level data from previous study, much more powerful, but not always available



## Historical control (2)

---

- Vulnerable to bias (same as those in concurrent control) plus
- Changes in outcome over time may come from change in:
  - underlying patient populations
  - criteria for selecting patients
  - patient care and management peripheral to treatment
  - diagnostic or evaluating criteria
  - quality of data available
- Do everything possible to minimize these!



## Historical control (3)

---

- Tend to exaggerate the value of a new treatment
- Even historical controls from a previous trial in the same institution or organization may still be problematic
  - Pocock (1977, *Brit Med J*) In 19 studies where the same treatment was used in two consecutive trials, differences in survival ranged from 46 to 24 , with four differences being statistically significant
- Adjustment for patient selection may be made, but all other biases will remain



# Historical or concurrent control: Statistical adjustments (1)

---

- Traditional methods
  - Matching
  - Covariate adjustment
- Newer methods
  - Bayesian
  - Propensity score
- Sensitivity analysis – estimate effect of unknown or unmeasured confounders on outcome → provides supportive assurance



# Historical or concurrent control: Statistical adjustments (2)

---

- Bayesian, propensity score and covariate adjustment
  - Essentially adjusts for differences in measured covariates between new treatment and control when estimating treatment effect
- Covariate adjustment is limited in the number of variables can include and interpretation can be challenging



# Bayesian methods (1)

---

- Can do everything Frequentist methods do and more
- Fundamental difference – provides coherent method for learning from evidence as it accumulates – combine prior information with current information, e.g., historical with concurrent trial data
- Bayesian predictive probability also has many uses
- 1998 FDA CDRH opened door for Bayesian in medical device trials and currently working on Guidance document – draft available at:  
<http://www.fda.gov/cdrh/osb/guidance/1601.html>
- Critical Path Opportunities opens door for drugs and biologics



## Bayesian methods (2)

---

- Control group usage
  - Non-randomized trial – use patient-level data from previous trial with Bayesian to control covariates
  - RCT – unbalanced enrollment borrow data from historical study to increase effective sample size of control group → reduce overall size
- Consult with Bayesian statistician – write me if need referrals
- Interesting FDA workshop on use of Bayesian in clinical trials available at:

<http://webcasts.prous.com/bayesian2004/index.asp>



# Propensity score methodology (1)

- Replace the collection of confounding covariates with one scalar function of these covariates: the propensity score  
**1 composite covariate:**



**Propensity Score**

**Balancing score**

- Excellent reference: *D'Agostino, R. Tutorial in biostatistics: propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group, Statist. Med. 17, 2265-2281, 1998*



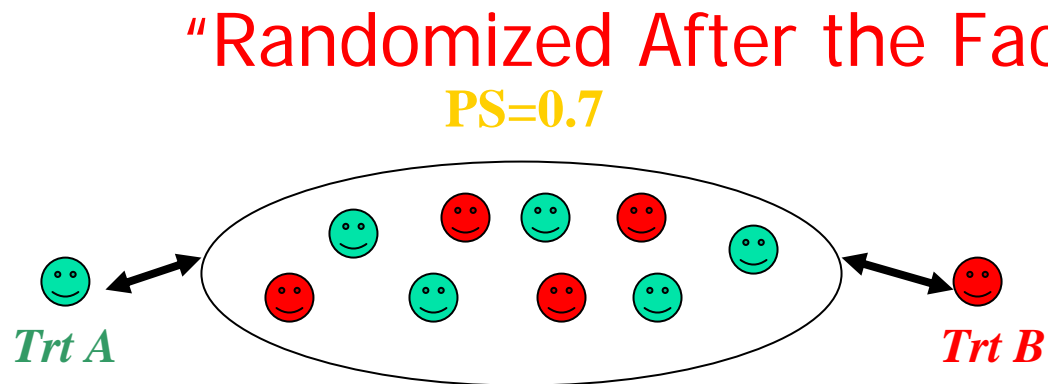
## Propensity score methodology (2)

---

- Propensity score (PS): conditional prob. of receiving treatment A rather than treatment B, given a collection of observed covariates
- Purpose: simultaneously balance many covariates in the two treatment groups and thus reduce the bias
- PS construction: multiple logistic regression model based on patient data of all measured covariates and actual treatment received

# Properties of propensity scores

- A group of patients with the same PS are equally likely to have been assigned to trt A
- Within a group of patients with the same PS, e.g., 0.7, some patients actually got trt A and some got trt B, just as they had been randomly allocated to whichever trt they actually received



L. Yue, FDA/Industry Stat Workshop 2005

IIR Clinical Trial Design 9/14/06



# Properties of propensity scores

---

- When the propensity scores are balanced across two treatment groups, the distribution of all the covariates are balanced in expectation across the two groups
- Use the propensity scores as a diagnostic tool to measure treatment group comparability
- If the two treatment groups overlap well enough in terms of the propensity scores, we compare the two treatment groups adjusting for the PS



# Stratification



- All patients are sorted by propensity scores
- Divide into equal-sized subclasses – usually 5
- Compare two trts within each subclass, as in a randomized trial; then estimate overall trt effect as weighted average
- It is intended to use all patients
- But, if trial size is small, some subclass may contain patients from only one treatment group



# Innovative control groups

---

- FDA Critical Path and other initiatives authorize us to find new ways to design trials and evaluate the results
  - #47 specifically works with virtual control groups consisting of databases, models and/or imaging collections
  - #53 record data on rare disease patients to create disease models and even virtual historical control groups



# Virtual control groups

---

- Not feasible or ethical to randomize
- Have good patient-level data on control of interest
- Create standard database of control group or model response statistically
- Compare to prospective study of new treatment(s) with adjustments
- Variation on historical control



# Redesigning epilepsy clinical trials

## Circumventing the ethics of placebo control

---

- Epilepsy, MS, Parkinson's disease, or other neurologic diseases with potential to cause harm if patients go untreated → once effective treatments exist, active placebo is no longer ethical option
- Past low-dose comparison trials with same exit criteria have provide very consistent results → people withdrawn from antiepileptic drug are 70% to 90% likely to exit
- From these data create historical control or virtual placebo
- New drug trial without placebo only needs to show lower exit rate with acceptable statistics, e.g., confidence interval



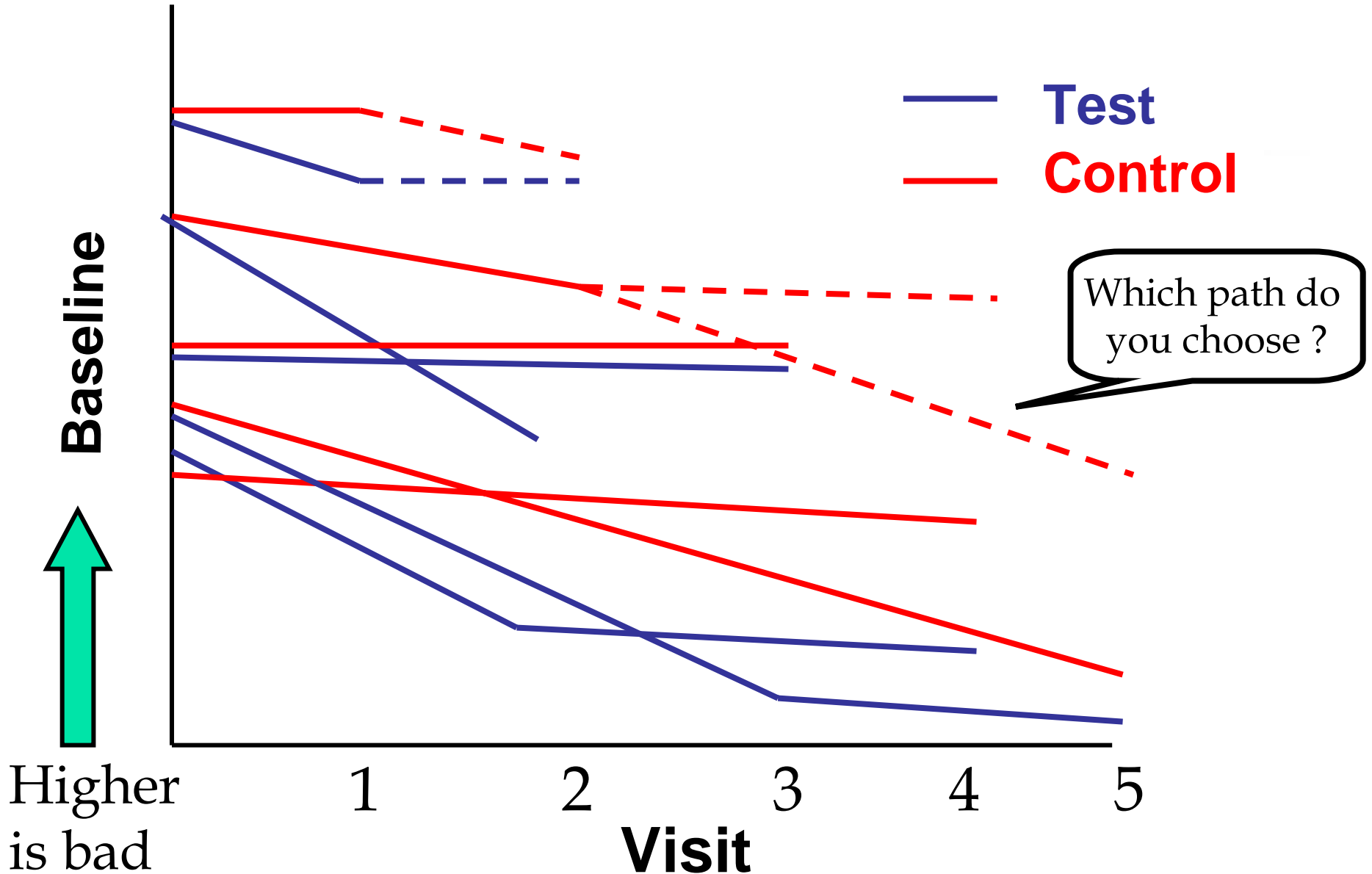
# Imputing missing data

---

- Drug becomes available before end of trial and patients switch to active – no more placebo data
- Measurements after switch are missing
- Develop Bayesian model for disease progression using historical data
- Incorporate information learned from the historical patients into similar model for placebo controls
- Use this model and on-protocol data to impute missing values

# What would be observed if subjects had stayed in trial ?

Impute values from subjects staying in longer and/or from dz progression model





# Alternative control groups: Summary

---

- FDA Critical Path and other initiatives have opened door to alternative control groups in drugs, biologics and devices
- E 10 Guidance provides many alternatives to placebo control
  - No treatment concurrent control
  - Dose-response concurrent control
  - Active (positive) concurrent control
  - External control (including historical control)
  - Multiple control groups
- Active control for non-inferiority requires use of proper study hypothesis and selection of appropriate reference control and NI delta
- External control study designs require minimizing bias through use of similar procedures and covariate adjustment
- Use innovative control groups in special situations where randomization is unethical and historical data is available



# Questions and Answers

---

*Thank you!*